APPENDIX B

Public Comments in Response to the *Federal Register* Notice (July 1, 2003)



May 2004

This Page Intentionally Left Blank



Advancing Science & Animal Institute for In Vitro Sciences, Inc.

August 14, 2003

Dr. William Stokes, D.V.M. NICEATM Director, NIEHS P.O. Box 12233, MD EC-17 Research Triangle Park, NC, 27709

Dear Dr. Stokes:

The members of the NICEATM and ICCVAM DCIWG committee are to be commended for the considerable effort that went into preparing the draft Minimum Performance Standards (MPS) documents for the three in vitro corrosivity assays. In requesting that these documents be prepared, the EPA has helped us all. Conceptually, the MPS documents are a substantial step forward in regulatory toxicology. They link the validation of an assay system (test system, protocol, endpoint determinations, controls, and prediction model) not only to the application of the assay system, but also to the production of data for regulatory review. These documents, and those that follow, will serve several purposes that are discussed in more detail below.

To start with some background, new test methods undergo several stages of maturation. A newly developed test will first be subject to prevalidation where the effectiveness of the technology transfer process and final protocol development occurs. The final protocol will be used to develop the prediction model that will allow the data from the new test to be calibrated against the desired toxicological action. A training set of reference test materials is used to develop the prediction model. Finally, the new test is subjected to formal validation, usually in several laboratories. For the validation study, a new set of reference test materials is employed. From the validation study, the performance characteristics of the new test (test system, protocol, and prediction model) are determined. Part of this process involves the identification of the essential elements of the test and test system; those elements (independent variables) that must be maintained/controlled to make the test reliable and predictive. This analysis should be performed with both proprietary and nonproprietary tests.

The ICCVAM submission guidelines require the use of controls (specifically positive controls). Performance norms for the positive control are established as part of acceptance criteria for a given "run" of the assay. The acceptable result obtained with the positive control helps to assure that the test system and test execution are functioning properly. While the concept of controls is not new to toxicology, the specification that the controls be performed concurrently with each unknown (or group of unknowns in a single batch) is new and tremendously important. The positive control provides a measure of consistency over time and across laboratories. The MPS documents also identify an important selection criterion for the positive control. The positive control must be able to demonstrate both over and under prediction (sensitivity) relative to the historical performance of the test. Using a 9-pound hammer (i.e., concentrated nitric acid) as a positive control is unlikely to effectively measure assay response. The discussion of benchmark controls (either chemical or formulation) is very helpful. While the positive

21 FIRSTFIELD ROAD SUITE 220

TEL: 301.947.6523

FAX: 301.947.6538

IIVS -- August 14, 2003

www.livs.org

control selected for a given assay should remain constant over time, benchmark materials tested concurrently with the unknowns would be selected to match the chemical class of the unknowns. The response of the benchmark controls facilitates interpretation of the results for the unknowns.

Validation studies are complex, time consuming and expensive. They serve to validate the complete test (test system [target tissue], protocol, endpoint measures, and prediction models). The successful validation of the test also tends to validate the mode of action measured by that test. For example, the mode of action for many corrosive chemicals is to penetrate the stratum corneum of the skin and rapidly kill the underlying keratinocytes. Conceptually, it is not hard to imagine modeling such a mode of action with an engineered human skin construct. However, modeling the quantitative (kinetic) aspects of the action is much more difficult. How much test material must be applied and for how long? How to measure the viability of the keratinocytes? How to translate the assay endpoint (e.g., percent viability) to a prediction of corrosive action? The test developer produces an assay protocol to address all of these parameters. The protocol may be based on a proprietary test system (e.g., skin construct) or assay endpoint (e.g., company X's ATP assay). Are those proprietary components of the test absolutely essential or could substantial equivalence be established for another test system or endpoint measure? By identifying the essential structural and functional elements of the test, the MPS approach will allow us (collectively) to draw on validation studies where a successful mode of action has been identified.

There are several additional reasons that the MPS approach is important:

- 1) In the original ECVAM-sponsored validation of in vitro assays for corrosivity, two skin constructs were tested. At the end of the validation program, neither skin construct was available commercially. Therefore, ZEBET conducted a study to show that the EpiDerm (MatTek, Ashland MA) construct was substantially equivalent to the validated tissue. Thus, the effort and expense of the validation study was not lost.
- 2) The Organization of Economic Cooperation and Development (OECD) prepares test guidelines for review and acceptance by its 30 member nations (including the United States). Their policy precludes specification of a proprietary test in OECD guidelines. As a result, the OECD has begun to specify structural and functional characteristics of a test (or test system) so that the guideline can draw on validation programs that employ proprietary methods or components.
- Some proprietary test developers may have made substantial investments in the validation programs for their test. Do the MPS guidelines diminish the economic value of that investment? We believe that they do not. The MPS guidelines provide a controlled mechanism for entry of a new test system or endpoint measure so that the field can grow, but they maintain and codify the standards for that assay. The MPS guidelines assume that the new or modified

"component" of the test will show substantial equivalence to that component of the validated test. For example, it is reasonable to expect that the substantially equivalent test will use the same prediction model as the validated test.

Otherwise, a new set of training test materials will be needed to develop the model. Clearly, one can not use the chemicals provided in the MPS to develop and then validate a prediction model! At some point in the number or degree of changes, a more complete validation of a modified method could be necessary.

Once a new test is accepted for regulatory use, additional laboratories are likely to begin using the method. The MPS documents provide the guidance needed to help demonstrate that the new test is being conducted properly. Successful execution of the test with the reference chemicals will help show that the equipment and reagents used in the new laboratory are within "normal limits" for the assay as it was validated. It will also help assess proper assay execution. The MPS guidelines are not a barrier to entry for a new laboratory but a means to link its performance with that of the validation laboratories. Data developed on unknown test materials would then be more credible for both the producers and users of such data.

Again, the authors of these documents are to be commended for developing the MPS concept and creating the subsequent documents. The format is well designed. I would ask however, that the authors become less prescriptive in their specifications for the report contents. Not every test substance will fit into the box that they have built. Perhaps more of the bullet points could include "if relevant to the conduct of the study". One item missing from the list is designation of the acceptance criteria (i.e., range acceptable positive control responses). For ease, the report section might have its own number (rather than being part of section 3).

The Minimum Performance Standards guidelines are an important step forward and NICEATM and ICCVAM DCIWG deserve a great deal of credit for their contribution.

Sincerely yours,

John W. Harbell, Ph.D. Chief Scientific Officer

Rodger D. Curren, Ph.D.

President



Roland Roguet
Business development EPISKIN
Department of reconstructed human skin models
L'ORÉAL RECHERCHE
90, rue du général Roguet
92583 Clichy Cedex France
tel +33 1 47 56 42 98
fax + 33 1 47 56 79 65

Comments on the draft « ICCVAM Minimum Performance Standards : In Vitro Human Skin Model Systems for Skin Corrosion »

The definition of minimum criteria of biological systems and performances on their uses is an essential guarantee to obtain relevant in vitro data and recognition of the method by regulatory authorities.

The initiative of the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Dermal Corrosivity and Irritation Working Group (DCIWG) corresponds to a real need for academics and industrial companies.

We have some suggestions in order to implement the draft.

The EPISKINTM model has been validated trough the ECVAM validation process and is on the market today. We suggest **to mention that EPISKINTM**, **besides EpiDermTM**, has a commercially available models for skin corrosivity assessment, page 4 line 20 and p7 line 17. In the same way we suggest to suppress the part of sentence " **and recommended as an alternative to EPISKINTM** page 3 line 17.

The ECVAM validation process demonstrated the ability of the EPISKIN™ test to discriminate corrosives from non-corrosives but also its ability to identify correctly known R35/I and R34II&III chemicals. This important advantage as regard to the UN packing groups classification should be mentionned in the draft. The addition of a sentence such as 'including the discrimination between R35 (UN packing group I) and R34 (UN packing groups II&III) after the citations (2,12,16) page 3 line 14 would complete the information provided on Table 1 concerning the corrosive subclasses recognized by some authorities.

CENTRE CHARLES ZVIAK
90, rue du Général Roguet-92583 Clichy Cedex FRANCE
Tél.: (33) 01 47 56 82 00 - Fax: (33) 01 47 56 75 27



Some taping errors have been found in the draft:

- page 3 line 18: the **citation 2** should be suppressed, since the ECVAM publication do not mention EpiDerm.
- page 4 line 21: citation number 22 instead of 221
- page 5 line 19: the exposure period of **4 hours** should be omitted since the draft described the method and predictive model for only two classes (corrosives/non-corrosives)(see prediction model page 7)
- page 6 line 9: add 'or' between glacial acetic acid and 8N KOH
- page 6 line 22, the word 'Cell' should be suppressed
- page 12: the title of the Table 3 should be corrected as "Accuracy of the Validated Human Skin Model EPISKIN™ Test Method for Skin Corrosion assessment"

We insist on the relevance of the draft proposed by the ICCVAM Dermal Corrosivity and Irritation Working Group and thank you for the opportunity to comment.

We hope the proposed modifications would help users of the method to better evaluate chemicals.

Sincerely,

Roland Roguet

August 14, 2003

Dr. William Stokes Director, NICEATM National Institute of Environmental Health Sciences P.O. Box 12233, MD EC-17 Research Triangle Park, NC 27709

Via electronic transmission to: iccvam@niehs.nih.gov

Dear Dr. Stokes:



HEADQUARTERS 501 FRONT STREET NORFOLK, VA 23510 TEL 757-622-PETA FAX 757-622-0457

These comments are submitted on behalf of People for the Ethical Treatment of Animals (PETA) and our more than 750,000 members and supporters in response to a July 1 notice in the *Federal Register* inviting public comment on three sets of "Minimum Performance Standards" for *in vitro* skin corrosivity tests proposed by the Dermal Corrosivity and Irritation Working Group of the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). We appreciate the work that has gone into the development of these documents and are hopeful that they will not only satisfy the needs of U.S. regulatory agencies, given their inability to lawfully require or recommend use of proprietary test methods, but will also be useful in preventing future bottlenecks in the validation pipeline both domestically and internationally.

PETA is in general agreement with the content of ICCVAM's proposed Minimum Performance Standards, with one notable exception: we strongly disagree with ICCVAM's recommendation that fully-validated *in vitro* human skin model systems (i.e., EpiDerm™ and EPISKIN™) be relegated to the status of merely "positive screens," whereby "substances that are negative *in vitro* might undergo additional testing in accordance with the tiered testing strategy" (*In Vitro* Human Skin Model MPS, p. 3), or, as articulated in ICCVAM's official recommendations to federal agencies: "Negative *in vitro* corrosivity responses shall be followed by *in vivo* dermal corrosion/irritation testing" (66 Fed. Reg. 49685).

As you know, both the European Union and the 30-member-country Organization for Economic Cooperation and Development (OECD) have accepted these validated *in vitro* human skin model systems either as stand-alone methods or as part of a purely *non-animal* weight-of-evidence strategy. Given ICCVAM's statutory mandate to promote the replacement, reduction, or refinement of animal-based testing and to strive for the elimination of unnecessary and duplicative efforts (42 *U.S.C.* Sec. 2851-3(b)), we cannot comprehend why ICCVAM persists in advocating a testing paradigm that is so clearly out-of-step with the international consensus on this issue.

It is also worth reiterating a point that was raised several times during the August 12-13 meeting of the National Toxicology Program's Scientific Advisory Committee on Alternative Toxicological Methods: that only a miniscule number (estimates range from two to six percent) of chemicals in commerce today are believed to possess irritating or corrosive properties. Thus, if regulatory agencies adhere to ICCVAM's testing recommendations (i.e., 66 Fed. Reg. 49685) and accept in vitro skin corrosivity assays as merely "positive screens," only a tiny handful of chemicals would likely be classified on the basis of in vitro data, while the overwhelming majority would still be required to undergo animal testing, ostensibly to "confirm" in vitro findings of non-corrosivity. From this perspective, ICCVAM's testing recommendations not only squander a golden opportunity for replacement, they promise to be equally meaningless and ineffectual from a reduction standpoint as well.

Even recognizing ICCVAM's stated concern regarding the potential for "false-negative" results *in vitro*, we should not need to remind the committee or its member agencies that the animal-based

Dr. William Stokes August 14, 2003 Page 2

reference data against which *in vitro* assays are so often compared have themselves seldom, if ever, been formally validated to demonstrate either their intra- or inter-laboratory reproducibility, much less their relevance to human beings. As just one example, we call your attention to a comparison of data from skin irritation tests on rabbits and skin patch tests on human volunteers for 65 substances, which found that nearly half—fully 45 percent—of classifications of chemical irritation potential based on animal tests were incorrect (MK Robinson et al. *Food Chem Toxicol* 40, 573-592, 2002).

As we have also pointed out in previous correspondence, a 1998 study by Worth and colleagues (ATLA 26, 709-720) determined that "false-negative" results from human skin equivalent models can be reduced to zero when combined with pH measurements and computerized structure-activity relationship modeling. The fact that this study is based on modeling data as opposed to a multi-chemical, multi-laboratory validation exercise should not, in itself, be seen to diminish the significance of the study's findings. Indeed, ICCVAM has already established a precedent for the acceptance of modeling data for validation purposes through its endorsement of the revised Up-and-Down Procedure for acute toxicity, the "validation" of which was based entirely on computer modeling.

Nonetheless, if ICCVAM and/or its constituent agencies had lingering doubts regarding the findings of Worth *et al.* (1998), they have had ample opportunity in the more than four years since this study was published to either confirm or refute its assertions. However, to the best of our knowledge, no such study has been undertaken by any ICCVAM member agency, which calls into question ICCVAM's continued resistance to a non-animal weight-of-evidence approach and its inexplicable insistence on "confirmatory" testing *in vivo*. Clearly, the former scenario is not only more humane, but also fully in harmony with the international consensus on this issue—both considerations being directly relevant to ICCVAM's statutory mandate.

With these considerations in mind, we strongly urge ICCVAM to revise its proposed Minimum Performance Standards and testing recommendations for *in vitro* human skin corrosivity systems to bring them into line with international regulations (e.g., EU Annex V) and testing guidelines (e.g., OECD 431).

Thank you for your attention and responsiveness to these comments.

Sincerely

Troy Seidle Science Policy Advisor

Federal Institute for Risk Assessment

BfR

Bundesinstitut für Risikobewertung

ZEBET at the BfR, Diedersdorfer Weg 1, 12277 Berlin, Germany

Dr. William Stokes
Director, NICEATM
NIEHS, MD EC-17
P.O. Box 12233
Research Triangle Park, NC, 27709
USA

zebet

Zentralstelle zur Erfassung und Bewertung von Ersatz- und Ergänzungsmethoden zum Tierversuch

Centre for Documentation and Evaluation of Alternatives to Animal Experiments

Diedersdorfer Weg 1 D-12277 Berlin Germany ☎ +49-1888-412-2270 Fax: +49-1888-412-2958 E-mail: zebet@bfr.bund.de P.O. Box 33 00 13 D-14191 Berlin Germany

T: (0)1888-412-0 Fax: (0)1888-412-4741 Internet: http://www.bfr.bund.de

your message date & initials

our document number & initials 7FRFT-2 / R4122 Phone +49-1888-412-2275

Berlin, 15 August 2003

Comments on ICCVAM Minimum Performance Standards on three types of *In Vitro* Tests for Skin Corrosion (Federal Register Notice Vol. 68, No. 126 / Tuesday, July 1, 2003, page 39104)

Dear Dr. Stokes

The institutions ZEBET and ECVAM have in 1997 already worked on the concept of a general use of skin models for regulatory toxicology. We have developed test protocols and prediction models that were generally applicable to different commercial skin models. For example, our skin model phototoxicity test developed with the full thickness skin model Skin_[Liebsch et al. Toxic. in Vitro 9, 557 – 562, 1994] could later be applied without any change to the epidermis model EpiDerm [Liebsch et al. Altex 14: 165 – 174, 1997], and was just recently successfully applied to the epidermis model SkinEthic [Jones et al. Toxic. In Vitro 17, 471-480, 2003]. Taking into account that experience and a comparable experience in the field of skin corrosion tests Michael Balls wrote in 1997 an ATLA editorial about definition of structural and performance criteria (copy enclosed) to facilitate the use of equivalent biological test systems in validated robust test methods. Finally, as you will recall, in the year 2002 we have internationally agreed on that concept in the OECD Workshop on Validation and Acceptance in Stockholm.

With this detailed introduction we want to emphasise that ZEBET very much welcomes the general concept and the definition of Minimum Performance Standards for the future use of "me too" test systems that claim to be equivalent to validated systems. In November 2001 this concept has been intensively discussed in the two OECD Extended Nominated Expert Consultations for the revision of Draft Test Guideline proposals on new Guidelines for Skin Corrosion and Phototoxicity, that finally resulted in accepted new OECD TG 430 and 431 on *Skin Corrosion*, and TG 432 on *Phototoxicity*. The Experts (incl. an ICCVAM representative) defined, for example, in TG 431 functional and performance criteria for new skin models in paragraphs 9, 10 and 11. In addition, 12 Reference Chemicals were defined that should be correctly classified if a new skin model was used or the test protocol modified. The Experts agreed that meeting these criteria is a sufficient proof of equivalency for a new skin model, and this was later confirmed by the National Co-ordinators of the OECD Member Countries. For TG 430 (TER Test), the same Reference Chemicals were defined to address the problem that the TER is sensitive to the rat strain used and the dimensions of the apparatus used. Here the twelve chemicals function as re-calibration chemicals rather than as a confirmation of the usability of the biological test system.

Because international consensus has been reached on OECD Test Guidelines 430 and 431, we welcome that the wording of these Guidelines has been used unchanged also in the ICCVAM MPS documents. However, ZEBET is opposing the additional mandatory requirement to test a

Berlin-Dahlem Thielallee 88 - 92 D-14195 Berlin Tel.: +49-1888-412-0 Berlin-Marienfelde Diedersdorfer Weg 1 D-12277 Berlin Tel.: +49-1888-412-0

Dessau Jahnstraße 8 D-06846 Dessau Tel.: +49-340-64000-0

2

larger set of chemicals with the TER and Skin Model Corrosion Test, since it results in mandatory re-validation of validated methods.

If testing a new skin model or a modified TER technology provides correct and reproducible results for the 12 OECD Reference Chemicals, then there is no need for testing additional chemicals, if we accept the robustness and general applicability of the new corrosion methods.

However, if not all of the 12 OECD Reference Chemicals are correctly classified additional refinement work and additional data is needed (depending on whether it looks promising). In that case, a list of well selected and easily available chemicals like the ones defined in the MPS documents can be very helpful. We therefore ask ICCVAM to accept the 12 OECD Reference Chemicals* and make it a mandatory requirement. The second set of 12 Test Chemicals should be recommended for test refinement when the 12 OECD Reference Chemicals have not 100% correctly been classified.

(* ICCVAM has deleted one of the twelve OECD Reference Chemicals (Acrylic Acid) from the list, because this was not included in the ECVAM Validation studies. However, the OECD experts had intentionally selected this chemical as a challenge for the skin model test, because it has a clear *in vivo* database as a strong corrosive.)

To emphasise our statement I can inform you that ZEBET and L'ORÉAL are currently very successfully co-operating on the generation of a common skin model test for *Skin Irritation Testing* that can be applied both to EPISKIN and EpiDerm models and that provides the same results in both models.

We do not comment in detail on the MPS document of the third Skin Corrosion Test (Barrier Test), since the situation is totally different: Because no OECD Test Guideline has been adopted, the ICCVAM MPS on the Barrier Test is not in conflict with international consensus. Moreover, to date the Barrier Method is still more a "black box" than the well validated and characterised skin models. Therefore, we support the definition of a sufficient number of reference chemicals, as suggested by the MPS document.

We do hope ICCVAM re-considers the TER and Skin Model MPS documents accordingly

On behalf of ZEBET

Sincerely yours

Dr. Manfred Liebsch

PS: We would like to put your attention to a few minor points (typos etc.):

Skin Model MPS:

Page 3, 3rd para: Although historically EpiDerm has been validated as an alternative to EPISKIN because it was not available any more, it was the catch up validation concept, only to show that EpiDerm was equivalent to EPISKIN. Delete that sentence, as EPISKIN is available again.

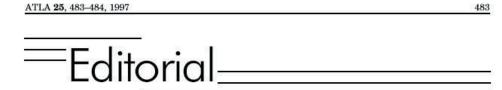
Page 4, 3rd para: Change reference (221) into (22)

Page 6, 4th para: Delete "cell"

Page 10, Table 2: As a strong MTT reducer that accumulates in the tissues n-Heptylamine is now correctly classified in all skin models (including SkiEthic), if the killed tissue control procedure is applied (see paragraph 15 of TG 431 and Liebsch et al ATLA 28, 371-401, 2000)

3

ENCLOSURE



Defined Structural and Performance Criteria would Facilitate the Validation and Acceptance of Alternative Test Procedures

The developers of new test procedures tend to want them to be tightly defined, so that they can gain their specific acceptance in the face of real or imagined competition, either for commercial reasons or to ensure that they gain the personal recognition they may deserve. However, it has become clear that this attitude is not in the interests of *in vitro* toxicology in general and may delay, or even prevent, the acceptance and application of scientifically relevant and reliable new approaches.

Three examples will illustrate the point. Firstly, Advanced Tissue Sciences withdrew their reconstituted human skin product, Skin^{2TM}, from the market, after it had been accepted by the US Department of Transport as a basis for classifying chemicals in terms of their skin corrosivity. Secondly, the withdrawal of Skin² and of EPISKINTM, a similar product made by Imedex, took place during a formal international study on in vitro tests for skin corrosivity, funded by ECVAM. Thirdly, Skin² was also in the process of being evaluated in the EU/COLIPA international validation study on in vitro tests for photoirritancy. As in the case of the withdrawal of a human skin product by Organogenesis a few years earlier, these developments led to annoyance and frustration since, whatever the manufacturers themselves had invested, and while one must sympathise with them, many other companies and laboratories had themselves invested considerable time and effort in evaluating the use of these systems for their own particular purposes. The results they had obtained had been most encouraging, which added to their sense of frustration.

This kind of problem could be avoided if, rather than validating and accepting particular kinds of commercial products, or methods involving particular cell lines, endpoints or endpoint assays, clearly laid down structural and performance criteria were to be defined and agreed for test systems to be used for particular purposes, then themselves subjected to prevalidation and formal validation. Any new test system which could meet these criteria would then be considered to be scientifically valid and acceptable, albeit after a small and independent confirmatory study in some circumstances.

It is for this reason that ECVAM and ZEBET are supporting studies on the applicability for *in vitro* corrosivity and photoirritancy testing of another human reconstituted human skin equivalent, EpiDermTM, made by MatTek, which, happily, promises to survive longer than its competitors. We are using our experience with Skin² and EPISKIN to speed up the acceptance of EpiDerm, not because we have any particular interest in MatTek or its products, but because we do not want much valuable experience to be wasted or the undoubted promise of this kind of test system to be lost.

At the same time, in order to provide one possible route of escape from the current impasse in the case of the acceptance of *in vitro* systems for percutaneous absorption, ECVAM has commissioned a study to define the structural and performance criteria which would be needed in such systems. Clearly, the structural characteristics required would include an effective barrier sufficiently similar to that found in the skin *in vivo*, and the performance criteria would include an ability to prevent the passage of certain standard test materials, while permitting the passage of others. Ideally, some of the *in vitro* systems should have the capacity to metabolise those kinds of test materials which would be likely to be metabolised by the human skin *in vivo*.

4

484 EDITORIAL

This having been done, ECVAM would be willing to support a prevalidation/validation study on *in vitro* systems which might meet the structural and performance criteria defined for percutaneous absorption testing.

This approach could be linked to the benchmarking concept as a possible route of escape from another impasse, namely, the absence of sufficient chemicals representative of the spectrum of chemicals to be tested in terms of type and scale of toxicity, backed by knowledge of sufficiently high quality. For example, an appropriate, and relatively small, set of standard materials which met *these* criteria, could be used to provide a standard curve, not only to establish the performance of the system on a particular occasion, but also as a means of expressing the result of the test on a novel material under investigation.

The structural and performance criterion approach could be taken further, since new tests could be developed to provide knowledge which is needed (i.e. to provide what Björn Ekwall has called "missing tests"). For example, it would be much more intelligent to devise realistic new tests for identifying human carcinogens, rather than merely speeding up the rodent bioassay or finding alternative methods for identifying chemicals which might be carcinogenic at high doses in rodents.

Michael Balls